JACC: CARDIOVASCULAR IMAGING VOL. ■, NO. ■, 2021

NEW RESEARCH PAPER

AI-Assisted Echocardiographic Prescreening of Heart Failure With Preserved Ejection Fraction on the Basis of Intrabeat Dynamics

Yu-An Chiou, BS,[a] Chung-Lieh Hung, MD,[b,c] Shien-Fong Lin, PhD[a,d]

ABSTRACT

OBJECTIVES The aim of this study was to establish a rapid prescreening tool for heart failure with preserved ejection fraction (HFpEF) by using artificial intelligence (AI) techniques to detect abnormal echocardiographic patterns in structure and function on the basis of intrabeat dynamic changes in the left ventricle and the left atrium.

BACKGROUND Although diagnostic criteria for HFpEF have been established, rapid and accurate assessment of HFpEF using echocardiography remains challenging and highly desirable.

METHODS In total, 1,041 patients with HFpEF and 1,263 asymptomatic individuals were included in the study. The participants' 4-chamber view images were extracted from the echocardiographic files and randomly separated into training, validation, and internal testing data sets. An external testing data set comprising 150 patients with symptomatic chronic obstructive pulmonary disease and 315 patients with HFpEF from another hospital was used for further model validation. The intrabeat dynamics of the geometric measures were extracted frame by frame from the image sequence to train the AI models.

RESULTS The accuracy, sensitivity, and specificity of the best AI model for detecting HFpEF were 0.91, 0.96, and 0.85, respectively. The model was further validated using an external testing data set, and the accuracy, sensitivity, and specificity became 0.85, 0.79, and 0.89, respectively. The area under the receiver-operating characteristic curve was used to evaluate model classification ability. The highest area under the curve in the internal testing data set and external testing data set was 0.95.

CONCLUSIONS The AI system developed in this study, incorporating the novel concept of intrabeat dynamics, is a rapid, time-saving, and accurate prescreening method to facilitate HFpEF diagnosis. In addition to the classification of diagnostic outcomes, such an approach can automatically generate valuable quantitative metrics to assist clinicians in the diagnosis of HFpEF. (J Am Coll Cardiol Img 2021;■:■–■) © 2021 The Authors. Published by Elsevier on behalf of the American College of Cardiology Foundation. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

From the [a]Department of Electrical and Computer Engineering, College of Electrical and Computer Engineering, National Chiao-Tung University, Hsinchu, Taiwan; [b]Division of Cardiology, Departments of Internal Medicine, MacKay Memorial Hospital, Taipei, Taiwan; [c]Institute of Biomedical Sciences, Mackay Medical College, New Taipei, Taiwan; and the [d]Institute of Biomedical Engineering, College of Electrical and Computer Engineering, National Chiao-Tung University, Hsinchu, Taiwan.
The authors attest they are in compliance with human studies committees and animal welfare regulations of the authors' institutions and Food and Drug Administration guidelines, including patient consent where appropriate. For more information, visit the Author Center.

Manuscript received October 7, 2020; revised manuscript received April 29, 2021, accepted May 6, 2021.

ISSN 1936-878X https://doi.org/10.1016/j.jcmg.2021.05.005

ABBREVIATIONS AND ACRONYMS

1D = 1-dimensional

A4C = apical 4-chamber

AI = artificial intelligence

BNP = brain natriuretic peptide

CNN = convolutional neural network

COPD = chronic obstructive pulmonary disease

GLS = global longitudinal strain

HF = heart failure

HFpEF = heart failure with preserved ejection fraction

LV = left ventricular

LVEF = left ventricular ejection fraction

ROC = receiver-operating characteristic

Heart failure (HF) is a significant emerging global public health issue associated with a high burden (approximately 1 million annually) of hospitalizations (1,2). Patients with HF can be classified into 2 primary phenotypes: those with HF with reduced ejection fraction and those with HF with preserved ejection fraction (HFpEF). Patients with HF with reduced ejection fraction have impaired systolic function, whereas those with HFpEF have diastolic dysfunction. These 2 phenotypes demonstrate distinct clinical and morphometric characteristics; therefore, they are assumed to derive from different pathologic molecular or biological mechanisms (3).

Ejection fraction is a percentage measurement of the amount of blood pumped out of the heart during systole. The most common cardiac systolic function assessment is through measuring the left ventricular ejection fraction (LVEF). The LVEF in patients with HF with reduced ejection fraction is <40%. In contrast, patients with HFpEF generally have LVEFs of more than 50%; thus, HFpEF cannot be determined solely from LVEF (4). To diagnose HFpEF is not an easy task; the diagnostic criteria need to follow the 2016 guidelines of the American College of Cardiology and the American Heart Association, take into account the diagnostic algorithm of the European Society of Cardiology, and include an LVEF ≥50%, LV mass index ≥115 mL/m^2 in men and ≥95 mL/m^2 in women, left atrial volume index ≥34 mL/m^2, along with functional alterations of lower tissue Doppler-derived myocardial relaxation (e′) and elevated left ventricular (LV) passive filling (E/e′ ratio), and elevated brain natriuretic peptide (BNP) concentration (5). HFpEF is gradually becoming the main form of HF. However, it usually takes about 10 min-30 min to perform a medical ultrasound examination and to process results from ultrasound tests. The accurate diagnosis of HFpEF depends on the clinician's expertise. Moreover, these ultrasound tests are performed only when the patient shows symptoms of HF.

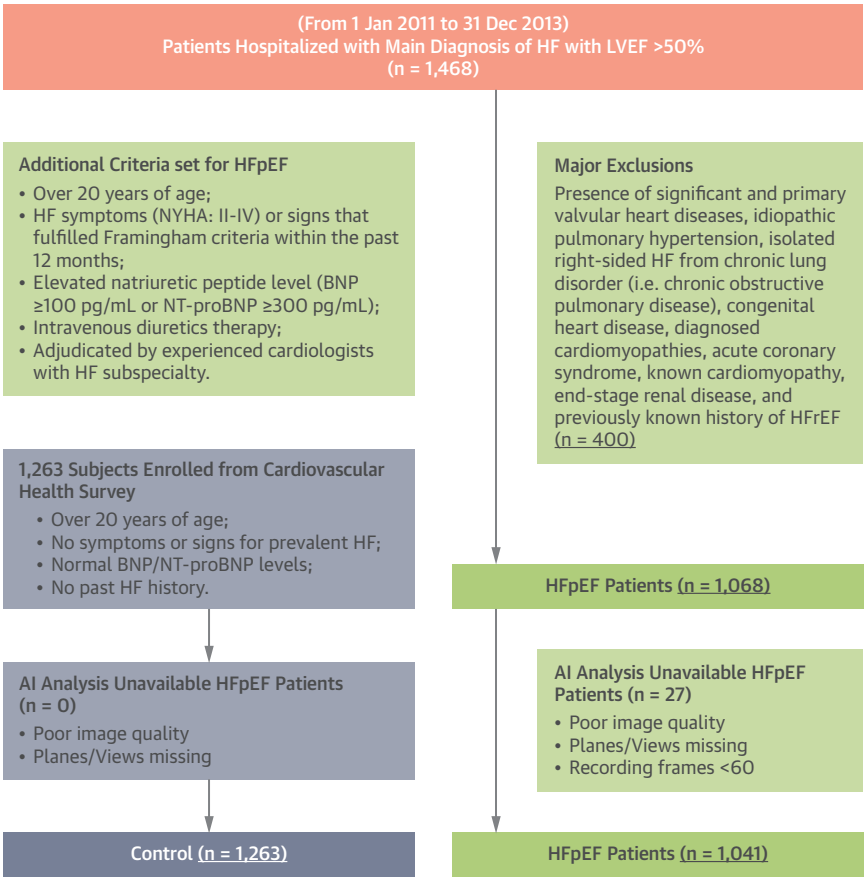
In recent years, artificial intelligence (AI) deep-learning algorithms have been widely applied in the medical field to assist in diagnosis (6). In particular, convolutional neural networks (CNNs) have been used in echocardiography and cardiac magnetic resonance to automatically extract cardiac functional parameters (7,8). AI-supported echocardiographic interpretation can accelerate the diagnosis of specific diseases (8). However, the diagnosis of diseases such as HFpEF, which requires multiangle echocardiography and measurement of multiple cardiac function parameters, remains challenging. HFpEF is also difficult to diagnose for inexperienced clinicians. Thus, the need for an efficient AI-assisted diagnostic prescreening tool for HFpEF is urgent. Previous studies assessed the differences in left atrial and LV global longitudinal strain (GLS) between participants with HFpEF and control subjects (9–12). The aim of this study was to design and evaluate a GLS-based AI prescreening system to assist in HFpEF diagnosis.

METHODS

STUDY DESIGN. We hypothesized that HFpEF could be identified by tracking the intrabeat dynamic changes in left atrial and LV regions in apical 4-chamber (A4C) view echocardiography. A data set comprising 1,263 asymptomatic individuals recruited from previous cardiovascular health surveys as the control group and 1,468 patients with HFpEF older than 20 years of age as the HFpEF group. The participants' 4-chamber view images were obtained from the echocardiographic files and randomly separated into training, validation, and internal testing data sets. The intrabeat dynamics of the geometric measures were extracted frame by frame from the image sequence to train the AI models. For further validation of the AI model, an external testing data set comprising 150 patients with symptomatic chronic obstructive pulmonary disease (COPD) and 315 patients with HFpEF from a different hospital. Accuracy, sensitivity, specificity, F1 score, and receiver-operating characteristic (ROC) curves were used to evaluate the prescreening outcome of the AI models.

DATASET SELECTION. This study was approved by the Mackay Memorial Hospital Institute Research Medical Ethics Committee (18MMHIS127e). The picture archiving and communication system at Mackay Memorial Hospital began in June 2009, when tissue Doppler information was available, and the study cohort of patients with HFpEF enrolled in the present AI analysis spanned January 2011 to December 2013. As HFpEF is a highly heterogeneous disorder, its clinical diagnosis can be challenging. The diagnosis of HFpEF in our retrospective cohort was established mainly by HF admission history (within the past 12 months) for symptoms or signs that fulfilled Framingham criteria and required intravenous diuretic therapy. The HFpEF diagnostic criteria required elevated natriuretic peptide level (BNP ≥100 pg/mL or N-terminal pro-BNP ≥300 pg/mL) (9), which was commonly used as a robust clinical index of such disease phenotype (such as in the GWTG-HF [Get

FIGURE 1 Flowchart of the Selection Process for the Internal Dataset



The internal data set included 1,263 asymptomatic individuals and 1,468 patients with heart failure with preserved ejection fraction (HFpEF). After data selection and exclusion, the final internal data set included 1,263 individuals without HFpEF with normal left ventricular ejection fractions (LVEFs) (>50%) as the control group and 1,041 patients with HFpEF. BNP = brain natriuretic peptide; HF = heart failure; HFrEF = heart failure with reduced ejection fraction; Nt-proBNP = N-terminal pro-brain natriuretic peptide; NYHA = New York Heart Association functional class.

With the Guidelines-Heart Failure] [10] and TOPCAT [Aldosterone Antagonist Therapy for Adults With Heart Failure and Preserved Systolic Function] trials [11]).

Our internal data set comprised 1,263 asymptomatic individuals recruited from cardiovascular health surveys as the control group and 1,468 patients with HFpEF older than 20 years of age as the HFpEF group (12). After data selection and exclusion, the final internal data set included 1,263 individuals without HFpEF with normal LVEFs (>50%) as the control group and 1,041 patients with HFpEF (total N = 2,304) (**Figure 1**). All the participants' Digital Imaging and Communications in Medicine files contained all angle views obtained on echocardiography. The study population characteristics of our internal data set, including echocardiographic parameters, are presented in **Table 1**. We performed additional validation using an independent external data set with 315 patients with HFpEF from another hospital (MacKay Memorial Hospital, Tamshui Branch) and 150 symptomatic patients with COPD. These patients with COPD (confirmed by spirometry) showed normal LVEFs (>50%) and low natriuretic peptide levels (N-terminal pro-BNP <125 pg/mL or BNP <35 pg/mL) without prior HF diagnosis or HF hospitalization.

ASSESSMENT OF DIASTOLIC FUNCTION. Conventional cardiac structural and geometric information was obtained according to contemporary guidelines (1). Assessment for each diastolic index and abnormal

TABLE 1 Study Population Characteristics

	HFpEF Group	Control Group
Physical information		
Individual amount	1,041	1,263
Age (yrs)	72 ± 14	47 ± 9
Male	489	346
Female	554	917
BMI (kg/m^2)	25 ± 5	24 ± 3
Systolic blood pressure (mm Hg)	149 ± 28	124 ± 16
Heart rates (beats/min)	80 ± 20	73 ± 10
NT-proBNP (pg/mL)	787 (290.3-1,780.0)	27.4 (11.7-49.9)
Total cholesterol (mg/dL)	170.72 ± 49.82	201.68 ± 34.37
High density lipoprotein cholesterol (mg/dL)	45.08 ± 13.71	53.52 ± 14.57
eGFR (mL/min/1.73 m^2)	47.40 ± 35.32	87.42 ± 15.81
Hypertension	706	0
Diabetes	467	0
Cardiovascular disease	477	1
Coronary artery disease	381	0
Atrial fibrillation	327	0
Echocardiographic information		
Interventricular septum (mm)	9.82 ± 1.84	9.18 ± 1.32
LV internal diameter in diastole (mm)	47.01 ± 5.63	47.68 ± 3.23
End-diastolic volume (mL)	102.08 ± 29.95	106.48 ± 16.18
End-systolic volume (mL)	42.18 ± 20.34	35.20 ± 7.73
LVEF (%)	60.24 ± 9.15	67.03 ± 4.33
LV mass index (g/m^2)	90.36 ± 27.79	77.65 ± 16.45
Deceleration time (ms)	209.99 ± 74.54	199.91 ± 31.66
TRV (m/s)	3.2 ± 0.5	2.1 ± 0.3
Average e′ (cm/s)	6.0 ± 2.3	11.7 ± 2.2
E/e′	16.8 ± 6.6	6.1 ± 1.6
LAVI (mL/m^2)	34.1 ± 12.2	15.6 ± 5.6

Values are n, mean ± SD, or median (interquartile range).

BMI = body mass index; E/e′ = ratio of mitral peak E velocity to average e′; eGFR = estimated glomerular filtration rate; HFpEF = heart failure with preserved ejection fraction; LAVI = left atrial volume index; LV = left ventricular; LVEF = left ventricular ejection fraction; NT-proBNP = N-terminal pro-brain natriuretic peptide; TRV = tricuspid regurgitation velocity.

cutoffs (including septal e′ <7 cm/s or lateral e′ <10 cm/s, left atrial volume index >34 mL/m^2, tricuspid regurgitation velocity >2.8 m/s, and ratio of mitral peak E velocity to average e′ [E/e′] >2.8m/s) were defined according to 2016 American Society of Echocardiography guidelines (13).

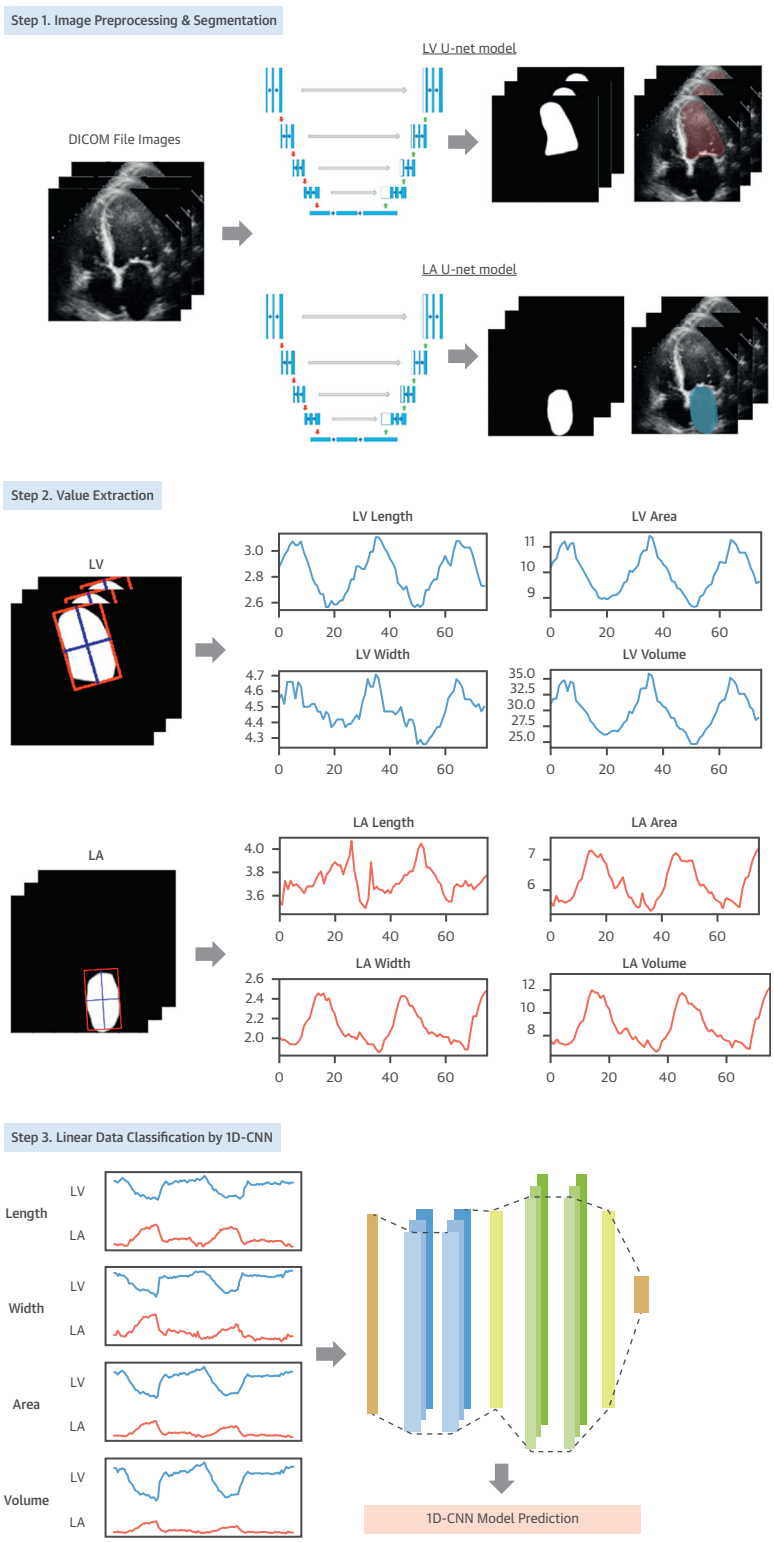
IMAGE PROCESSING. The step-by-step image analysis procedure is presented in Figure 2. All A4C view frames were extracted from the Digital Imaging and Communications in Medicine files and converted into images, which were cropped and resized to 128 × 128 pixels. Each pixel was then normalized between 0 and 1 for neural network training and prediction. Edges of the left atrial and LV chambers were obtained using U-net, a conventional neural network with a U-shaped architecture consisting of a contracting and an expanding pathway focused on image features (14). The U-net depth applied in this research was 4 layers for down-convolution and another 4 layers for up-convolution. This U-net structure, which allows the classification of each pixel, is widely used in automatic biomedical image segmentation (15,16). During U-net processing, each A4C view image of the left atrial and LV regions was captured to calculate chamber length, width, area, volume, and LVEF. For U-net training, a total of 500 A4C view images from internal data set were selected, including 450 A4C view images for the training data set, and 50 randomly selected untrained A4C view images served as the testing data set (Figure 3, left).

In the value extraction process, the length, width, area, and volume were extracted as intrabeat dynamics. After U-net segmentation, the length, width, area, and volume of the LV and left atrial information were calculated using a convex hull algorithm, which was used to find the smallest square around the edge of the chamber. All parameters in each frame were recorded as linear data for the following 1-dimensional (1D) CNN processing. The maximum and minimum areas of the left atrium and left ventricle were defined as the end of a diastolic view and the end of a systolic view. Calculation of left atrial and LV volumes and derived LVEF was based on recommended guidelines for cardiac chamber quantification by echocardiography (1).

DATA CLASSIFICATION BY 1D CNN. One-dimensional CNNs are extensively used to classify medical signals (17). Convolutional and pooling layers are used to concentrate 1D linear signals to discover certain disease characteristics. The intrabeat dynamic changes of left atrial and LV length, width, area, and volume were recorded as linear waveform signals, which can be trained and further classified by a 1D CNN. These 4 types of linear signals were cropped to 60 frames, which constituted at least 1 complete systolic-diastolic cycle, and were then separately imported into 4 identical 1D CNN models. The network structure consisted of 4 convolutional layers, 2 pooling layers, 2 dropout layers, and 1 dense layer. The convolutional and pooling layers were used to concentrate the linear features. The dropout layers were designed to randomly drop some connections between layers to prevent overfitting, and the dense layer was designed for output.

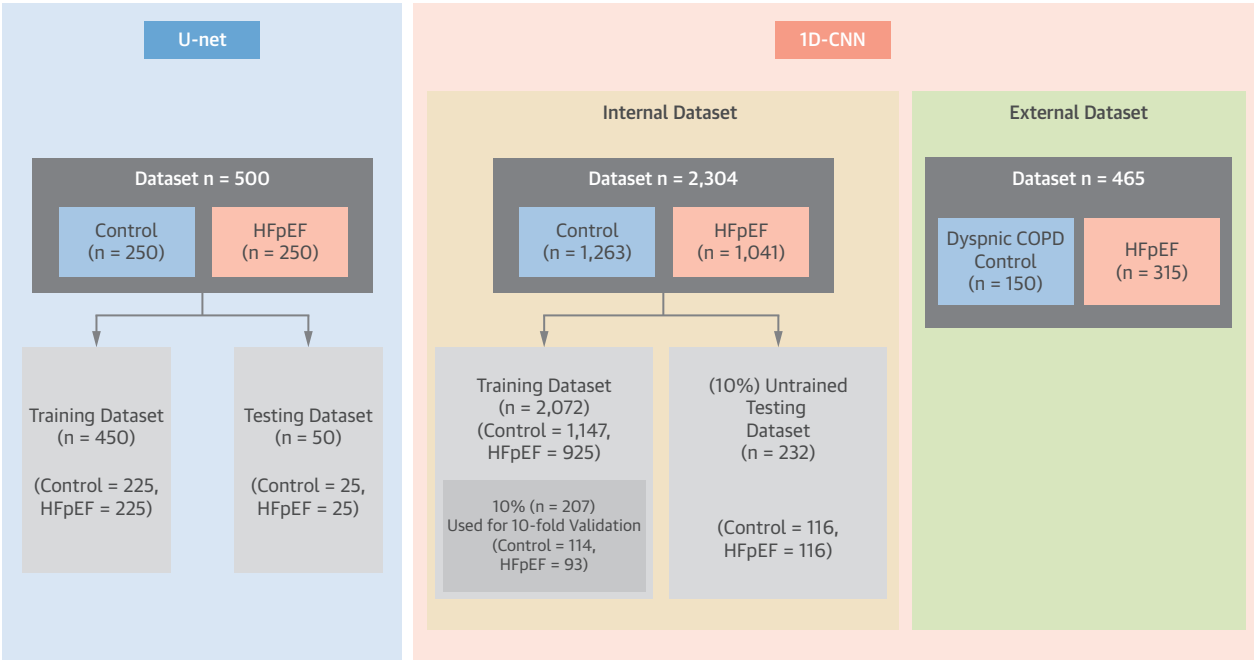
STATISTICAL ANALYSIS. For U-net model evaluation, the mean intersection over union, a common evaluation metric for semantic image segmentation, was applied. It can be used to count the number of overlapping pixels between the AI neural network prediction and manual labeling areas (8). The mean intersection over union formula is as follows:

FIGURE 2 **Step-By-Step Procedures for Image Preprocessing and Segmentation, Value Extraction, and 1D CNN Model Prediction**



Each dynamic left ventricular (LV) and left atrial (LA) length, width, volume, and area data (60 frames) were extracted from the U-net prediction images. These linear signals can be trained for model prediction by a 1-dimensional (1D) convolutional neural network (CNN). DICOM = Digital Imaging and Communications in Medicine.

FIGURE 3 Training and Testing Dataset Used for U-Net and 1D CNN Model



The internal data set for 1D CNN comprised patients with HFpEF and asymptomatic individuals, with the external data set further validated to distinguish patients with HFpEF from those with dyspnea and chronic obstructive pulmonary disease (COPD). Abbreviations as in Figures 1 and 2.

$$\text{Mean intersection over union} = \frac{\text{Manual label} \cap \text{prediction}}{\text{Manual label} \cup \text{prediction}}$$

Furthermore, the correlation between the ground truth label and the AI prediction results can be assessed using the Spearman correlation coefficient and a Bland-Altman plot. The Spearman correlation coefficient (ρ) represents the level of correlation, and a P value of <0.01 indicates a significant correlation. Bland-Altman plots are used to analyze the agreement between 2 different assays. Both methods are commonly used in the medical field to compare the differences between the ground truth label and AI prediction results (8).

For evaluation of the 1D CNN model, accuracy, sensitivity, specificity, F1 score, and a ROC curve, commonly applied in the evaluation of deep-learning models, were used. ROC curves illustrate the diagnostic capabilities of a classifier system for model prediction (eg, HFpEF or non-HFpEF). The area under the ROC curve is equal to the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one (18). For the F1 score, both precision and recall are considered to compute the score to evaluate the classifier (19). Cohen's kappa coefficient (κ), a statistic that measures inter-rater reliability for categorical items (20), was used to calculate the error between the model-predicted label and manual echocardiographic label. In this study, the kappa score was used to evaluate the expert-calculated LVEF and the AI-derived LVEF to measure the inter-rater consistency between the expert and the AI model. In addition, the kappa score was also used to evaluate the HFpEF prediction results between the expert and the AI model.

RESULTS

The internal data set selection process is displayed in Figure 3 (right). To train our AI model, we randomly separated the internal data set (total N = 2,304) into a training data set (n = 2,072) and a 10% untrained internal testing data set (n = 232) (Table 2). Four types (including length, width, area, and volume) of left atrial and LV linear signals were used for our HFpEF model prediction. From the training data set, 10% were randomly selected during the training process

(as internal validation; n = 207) for 10-fold validation. As previously mentioned, our 1D CNN model was further validated in an external testing data set comprising independent patients with HFpEF and patients with symptomatic COPD without HF (Figure 3, right; Table 3). We further tested the diagnostic performance of our trained AI model in identifying each abnormal diastolic index (as a binary variable) proposed in contemporary guidelines both in internal and external testing data sets (17).

INTERNAL DATA SET. After U-net model segmentation, the left atrial and LV length, width, area, and volume of each segmented image as linear signals were used for 1D CNN classification. After training and validation from internal data set, we tested our 1D CNN model in the 10% randomly selected untrained internal testing data set (n = 232) (Table 2). The average time spent on each data set analysis was <3 min. The internal testing data set results revealed that compared with dynamic information on length, width, and volume, both left atrial and LV area changes showed the highest diagnostic performance (area under the curve [AUC]: 0.95 for both) in distinguishing patients with HFpEF from control subjects (Table 4). These scores represent the precision in classifying patients with HFpEF and those without HFpEF. Figure 4A presents the ROC curve results for the internal testing data set. The left atrial area model had the highest prediction accuracy among the left atrial models, with accuracy, sensitivity, specificity, F1 score, and kappa score of 0.91, 0.96, 0.85, 0.91, and 0.81, respectively. Similarly, the LV area model had the highest prediction accuracy among the LV models, with accuracy, sensitivity, specificity, F1 score, and kappa score of 0.88, 0.91, 0.85, 0.88, and 0.75, respectively. Overall, our AI model showed satisfactory performance in identifying each abnormal diastolic index from our internal testing data set (Supplemental Tables S1 to S4). Comparison of ground truth labels and AI predictions of end-diastolic volume, end-systolic volume, and LVEF from an untrained internal data set (n = 232), in which Spearman correlation coefficients and differences (Bland-Altman plot) were further detailed in Supplemental Figure S1. Supplemental Figure S2 shows the ROC curves.

EXTERNAL DATA SET. The external testing data set was used to verify that the AI model could fit patients with HFpEF from different hospitals. Patients with COPD without HFpEF were included in the external testing data set as a control group because the external symptoms of COPD are similar to those of HFpEF. The external testing data set comprised 150 patients with symptomatic COPD without HF and 315 patients with HFpEF independent of our internal data set (Table 3). The external cohort results showed that the linear signal data of both left atrial and LV area had higher accuracy than other parameters in predicting HFpEF (Table 5). Figure 4B presents the ROC curve results from the external testing data set, which showed a similar pattern to the original internal testing data set. These consistent results support that our AI-assisted prescreening system shows potential in distinguishing HFpEF from noncardiac causes of dyspnea among symptomatic patients with COPD. Additionally, our AI model also performed well in identifying each abnormal diastolic index from external testing data set (Supplemental Tables S5, S6, S5 to S8), with ROC curves shown in Supplemental Figure S2.

TABLE 2 Internal Testing Dataset Population Characteristics

	Internal Control Group	Internal HFpEF Group
Physical information		
Individual amount	116	116
Age (yrs)	48 ± 9	68 ± 8
Male	77	66
Female	39	50
BMI (kg/m^2)	23.± 3	27 ± 6
Systolic blood pressure (mm Hg)	124 ± 18	150 ± 33
Heart rates (beats/min)	72 ± 10	85 ± 17
NT-proBNP (pg/mL)	27.4 (11.8-64.4)	742 (176-2,000.0)
Total cholesterol (mg/dL)	201.18 ± 35.32	185.47 ± 58.78
High density lipoprotein cholesterol (mg/dL)	53.80 ± 16.21	40.28 ± 10.92
eGFR (ml/min/1.73 m^2)	89.42 ± 17.22	51.12 ± 32.54
Hypertension	14	68
Diabetes	3	50
Coronary artery disease	6	43
Atrial fibrillation	0	22
Echocardiographic information		
Interventricular septum (mm)	9.11 ± 1.42	9.78 ± 2.31
LV internal diameter in diastole (mm)	47.54 ± 3.80	48.92 ± 6.43
End-diastolic volume (mL)	104.77 ± 16.87	115.19 ± 35.05
End-systolic volume (mL)	34.32 ± 8.510	49.89 ± 24.55
LVEF (%)	67.42 ± 3.89	56.96 ± 11.31
LV mass index (g/m^2)	80.53 ± 20.10	89.22 ± 31.26
Deceleration time (ms)	204.45 ± 35.15	174.62 ± 52.54
TRV (m/s)	2.3 ± 0.3	3.2 ± 0.5
Average e′ (cm/s)	11.0 ± 2.2	5.8 ± 1.5
E/e′	6.6 ± 1.7	17.2 ± 5.5
LAVI (mL/m^2)	16.0 ± 5.8	35.7 ± 9.1

Values are n, mean ± SD, or median (interquartile range).
Abbreviations as in Table 1.

DISCUSSION

The AI-assisted prescreening system developed in this study can automatically extract featured

TABLE 3 External Testing Dataset Population Characteristics

	External COPD Control Group	External HFpEF Group
Physical information		
Individual amount	150	315
Age (yrs)	55.03 ± 13.43	70 ± 13
Male	131	134
Female	17	181
BMI (kg/m^2)	71.19 ± 12.44	25.± 5
Systolic blood pressure (mm Hg)	130.24 ± 19.70	152 ± 27
Heart rates (beats/min)	70 ± 10.8	84 ± 23
NYHA functional class		
I	–	13
II	–	190
III and IV	–	112
Cardiomegaly	16	167
Jugular vein engorgement	12	76
Peripheral oedema	5	192
NT-proBNP (pg/mL)	33.70 (16.00-93.42)	796 (399.50-2,000.00)
Total cholesterol (mg/dL)	224.94 ± 34.48	177 ± 55
High density lipoprotein cholesterol (mg/dL)	47.2 ± 10.7	45.56 ± 13.82
eGFR (mL/min/1.73 m^2)	67.5 ± 24.2	49.47 ± 31.73
Hypertension	48	162
Diabetes	35	160
Coronary artery disease	29	137
Atrial fibrillation	6	105
Echocardiographic information		
Interventricular septum (mm)	9.56 ± 1.48	10.72 ± 2.26
LV posterior wall thickness (mm)	9.58 ± 1.51	10.53 ± 1.83
LV internal diameter in diastole (mm)	50.60 ± 38.05	48.43 ± 6.38
End-diastolic volume (mL)	106.37 ± 20.12	112.53 ± 35.10
End-systolic volume (mL)	37.20 ± 8.90	48.90 ± 27.24
LVEF (%)	65.12 ± 5.41	56.90 ± 13.51
LV mass index (g/m^2)	82.65 ± 22.20	96.91 ± 30.44
Deceleration time (ms)	207.02 ± 48.36	207.69 ± 67.38
TRV (m/s)	2.4 ± 0.3	3.2 ± 0.5
Average e′ (cm/s)	10.0 ± 1.9	5.9 ± 1.8
E/e′	7.2 ± 2.0	16.9 ± 6.8
LAVI (mL/m^2)	21.7 ± 5.3	36.6 ± 12.4

Values are n, mean ± SD, or median (interquartile range).
NYHA = New York Heart Association; other abbreviations as Table 1.

intrabeat dynamic patterns associated with HFpEF to assist in diagnosis. Echocardiographic interpretation through AI assistance can be helpful to less experienced clinicians and reduce interobserver and intraobserver variability (21). Our system provides an automatic interpretation that exhibits high accuracy in detecting HFpEF. However, because of the limited amount of data available for patients with HFpEF, our system could only provide a proof of concept for AI model development to aid in HFpEF prescreening. To improve our system, further research and data collection are required.

Our study concept was based on previous research in which GLS was used as a predictor of HFpEF. HFpEF is divided into 4 stages of severity (A-D). In Stages B to D, the resting left atrial, LV, and right ventricular GLS of the patient decrease gradually with advancing stage (22). GLS < 16% has been set as a new minor criterion to assist in HFpEF diagnosis (5). To improve the diagnostic feasibility of GLS, instead of extracting only LV diastolic and LV systolic length for GLS calculation, our AI model incorporated dynamic changes (and hence waveform characteristics) of several left atrial and LV parameters (including length, width, area, and volume) during the systolic-diastolic cycle. The detailed intrabeat dynamic features were extracted using a 1D CNN for classification. Because the area formula contained both length and width changes, the intrabeat dynamic changes of LV and left atrial area had the highest accuracy among all parameters (0.88 and 0.91, respectively). Thus, changes in dynamic area rather than dynamic length may be more helpful in assisting in HFpEF diagnosis.

Specific features in dynamic LV area changes in the HFpEF group were manually classified on the basis of cardiac phase-specific analysis. **Figure 5** depicts a typical illustration of the dynamic LV area changes between the HFpEF and control groups. **Figure 6** demonstrates the chronologic differences in intrabeat dynamic LV area changes between the control group and patients with several subtypes of HFpEF. Intrabeat dynamics in the control group showed an early upstroke followed by a steady rise or plateau-like diastolic waveform, representing a sustained volume for approximately 15 frames-25 frames (0.3 s-0.5 s). In contrast, in the HFpEF group, 72% (including the round-peak, sharp-peak, 2-peak in systole, and irregular groups) of the patients with HFpEF showed no steady plateau phase during the diastolic stage. These may result in sharper waveforms during the diastolic than the systolic phase, which likely reflects markedly impaired diastolic mechanics (eg, reduced e′ and higher E/e′ indicative of elevated LV filling pressure) or delayed relaxation in HFpEF (**Figures 5 and 6**). This was partly supported by the findings that our AI model is also capable of differentiating abnormal diastolic indexes in both our internal and external testing data sets. The irregular group (10%) did not show a regular systolic-diastolic cycle. The fake control group (21%) demonstrated perfect dynamic area changes, as did the control group, in which clinicians could not find specific HFpEF characteristics. AI technology is thus needed to discover the subtle features hidden in such data.

PREVIOUS RESEARCH. Conventional GLS measurement requires clear LV apical views (eg, the A4C view) for speckle-tracking echocardiography, which is used

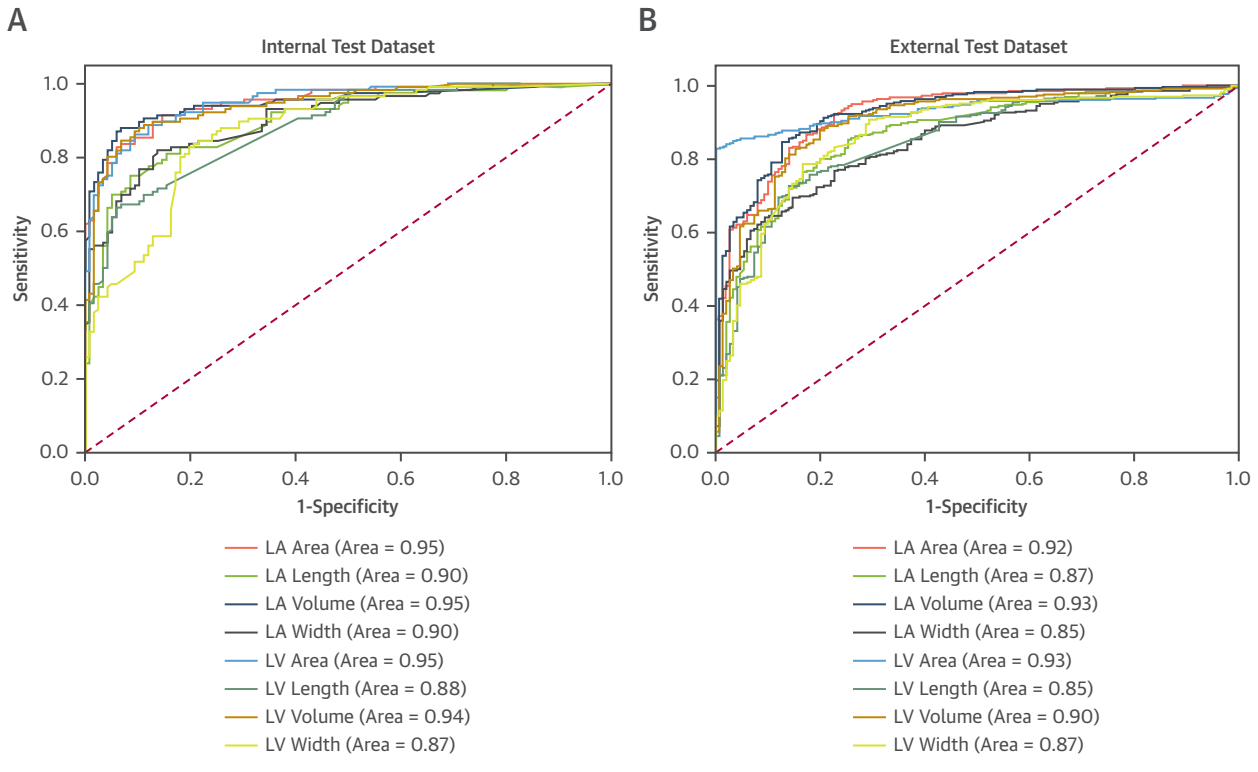
TABLE 4 Artificial Intelligence Prediction Results for the Internal Testing Dataset (n = 232, Including 116 Control Subjects and 116 Patients With Heart Failure With Preserved Ejection Fraction)

	U-Net Prediction Result	1D CNN Prediction Result					
	Mean IoU	Accuracy	Sensitivity	Specificity	F1 Score	κ	AUC (95% CI)
Left atrium							
Length	0.89	0.81	0.92	0.69	0.83	0.61	0.89 (0.85-0.93)
Width		0.83	0.89	0.77	0.84	0.65	0.90 (0.86-0.94)
Area		0.91	0.96	0.85	0.91	0.81	0.95 (0.92-0.97)
Volume		0.89	0.94	0.84	0.89	0.78	0.94 (0.91-0.97)
Left ventricle							
Length	0.90	0.78	0.9	0.67	0.81	0.56	0.88 (0.83-0.92)
Width		0.81	0.81	0.80	0.81	0.61	0.87 (0.82-0.91)
Area		0.88	0.91	0.85	0.88	0.75	0.95 (0.92-0.97)
Volume		0.80	0.82	0.78	0.85	0.68	0.94 (0.91-0.97)

1D = one-dimensional; AUC = area under the curve; CNN = convolutional neural network; IOU = intersection over union; LA = left atrial; LV = left ventricular.

to count HFpEF GLS ranges (22). However, GLS remains a minor criterion in the diagnosis of HFpEF for which no specific range has been set. Therefore, a variety of geometric or diastolic functional indexes may be required to establish the diagnosis of HFpEF (4-6,17,18). Furthermore, current echocardiographic indexes do not take dynamic information of these functional measures into account. In the present study, patterns of intrabeat dynamic changes of LV and left atrial area, length, width, and volume can be automatically determined and analyzed using AI during the entire systolic-diastolic cycle instead of single

FIGURE 4 ROC Curves for 1D CNN in HFpEF-Specific Prediction Model From Internal and External Testing Datasets



Receiver-operating characteristic (ROC) curves for the internal testing data set **(A)** and the external testing data set **(B)**. Abbreviations as in Figure 2.

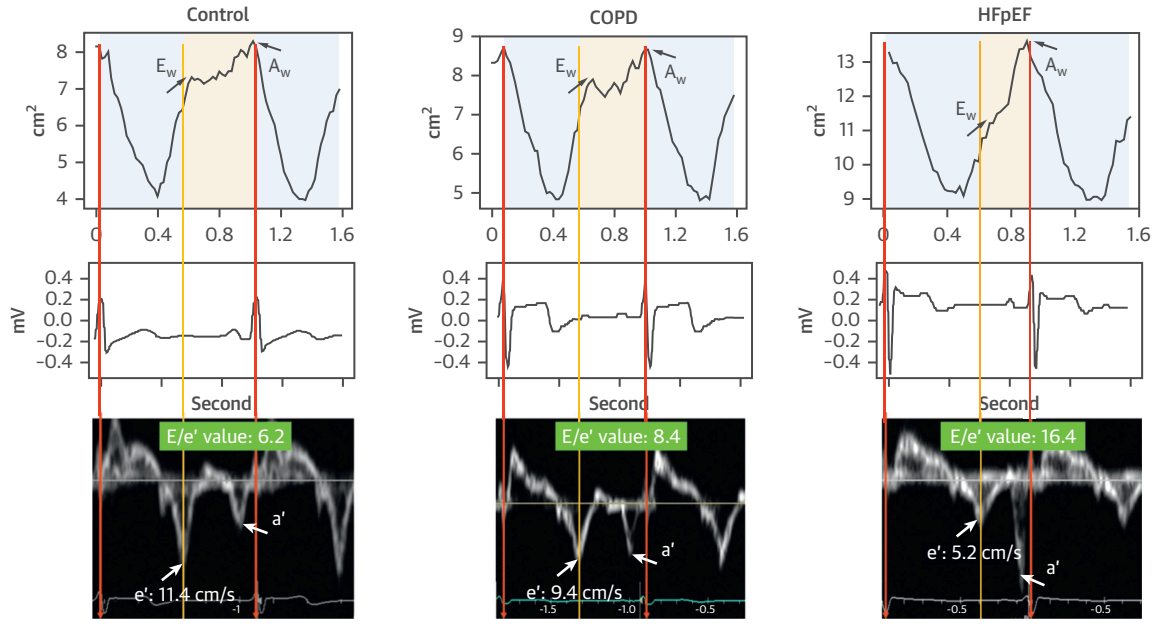
TABLE 5 Artificial Intelligence Prediction Results for the External Testing Dataset (n = 465, Including 150 Control Subjects With Chronic Obstructive Pulmonary Disease and 315 Patients With Heart Failure With Preserved Ejection Fraction)

	U-Net Prediction Result	1D CNN prediction result					
	Mean IoU	Accuracy	Sensitivity	Specificity	F1 Score	κ	AUC (95% CI)
Left atrium							
Length	0.88	0.80	0.75	0.83	0.71	0.57	0.87 (0.84-0.89)
Width		0.77	0.66	0.82	0.64	0.47	0.85 (0.82-0.88)
Area		0.85	0.83	0.85	0.77	0.66	0.92 (0.90-0.94)
Volume		0.86	0.85	0.86	0.79	0.69	0.93 (0.90-0.95)
Left ventricle							
Length	0.92	0.77	0.86	0.73	0.71	0.53	0.85 (0.82-0.88)
Width		0.84	0.71	0.90	0.74	0.62	0.87 (0.84-0.90)
Area		0.87	0.84	0.88	0.80	0.70	0.93 (0.91-0.95)
Volume		0.85	0.79	0.89	0.78	0.67	0.95 (0.88-0.93)

Abbreviations as in Table 4.

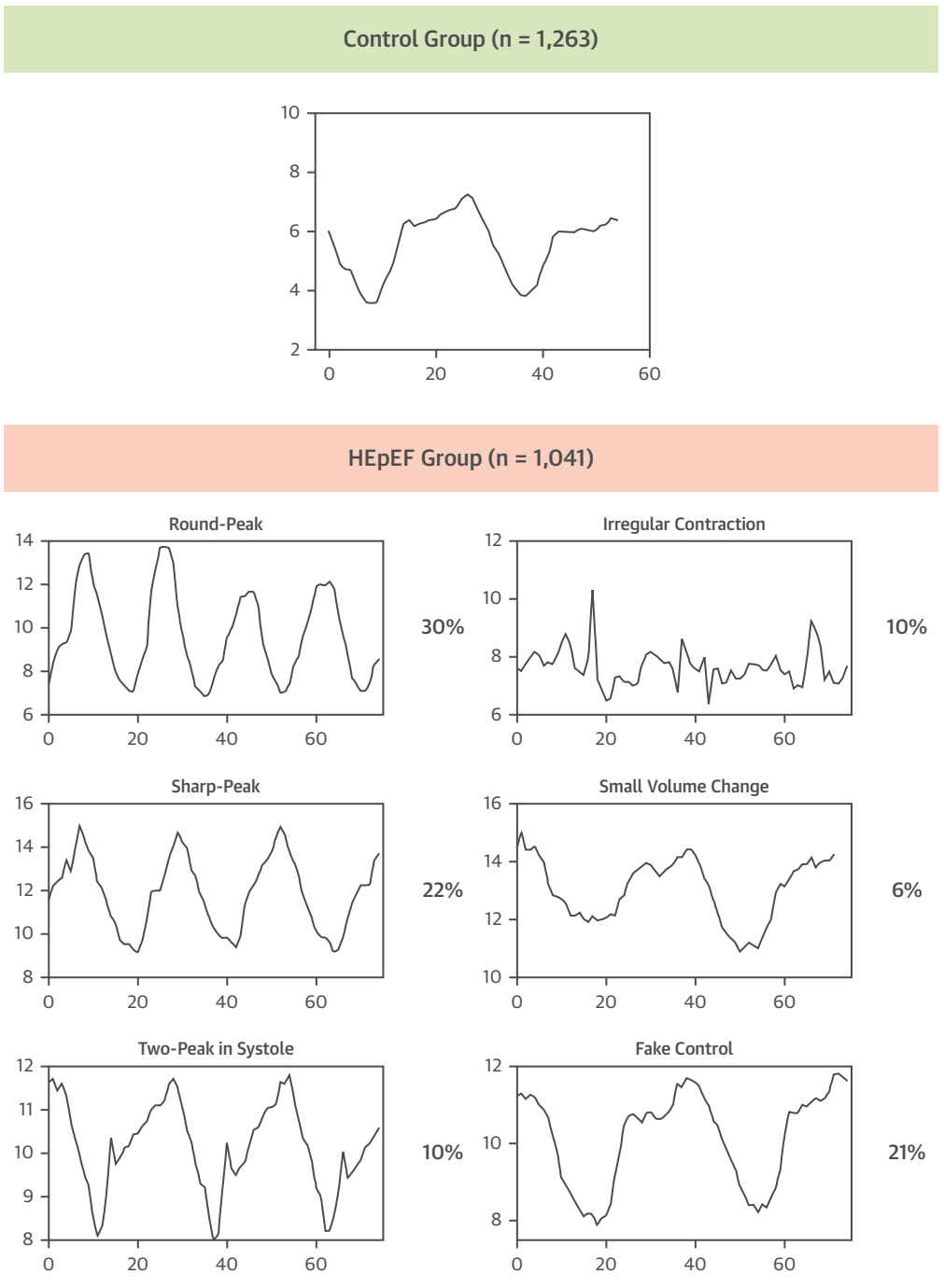
GLS value. This will probably generate more comprehensive information (including waveform characteristics on phasic changes, heart rate, or heart cycle regularity) than do single GLS values. These specific features, when integrated as linear signals and captured by 1D CNNs, could be supplemental and provide mechanical insights into altered left atrial and LV kinetics for patients with HFpEF compared with conventional echocardiographic parameters. Indeed, our 1D CNN model showed remarkable performance in HFpEF model prediction and discriminated impaired diastolic mechanics (17), which were further validated

FIGURE 5 Schematic Illustration of Typical Intrabeat Dynamic LV Area Changes in Asymptomatic Control Subjects, Patients With COPD, and Patients With HFpEF



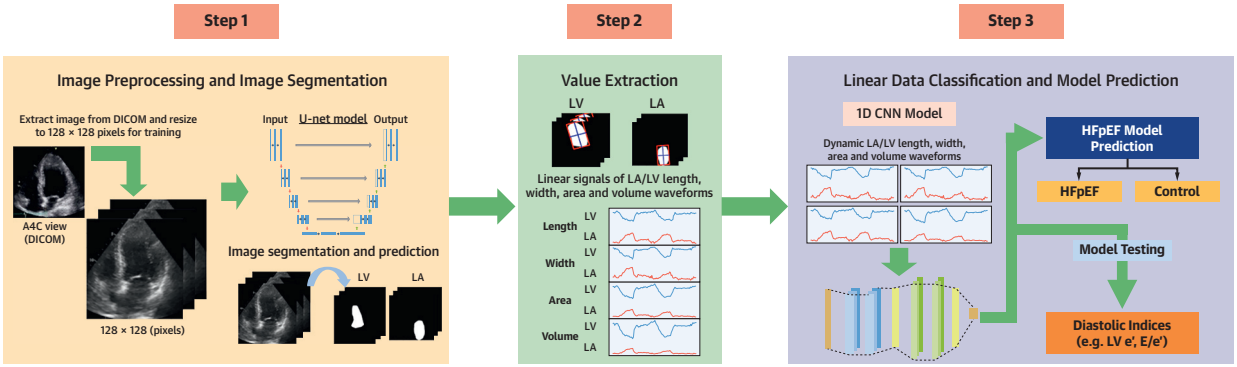
Chronologically matched cardiac phase-specific (gated using both electrocardiography and mechanical events) waveform showing impaired diastolic mechanics and delayed relaxation in HFpEF, which likely result in sharper diastolic waveform compared with control and COPD groups. **Red and yellow lines** indicate end-diastolic and peak e′ time point, respectively. Intrabeat dynamics Ew and Aw correspond to e′ and a′ time points, respectively. Abbreviations as in Figure 1 and 3.

FIGURE 6 **Intrabeat Dynamic Area Changes in the Control and HFpEF Subgroups**



Patients with heart failure with preserved ejection fraction (HFpEF) may exhibit different types of specific dynamic left ventricular area changes during systolic-diastolic cycles.

CENTRAL ILLUSTRATION Framework Diagram for Artificial Intelligence-Assisted Heart Failure With Preserved Ejection Fraction Prescreening Process



Chiou, Y.-A. et al. J Am Coll Cardiol Img. 2021;■(■):■-■.

The heart failure with preserved ejection fraction prescreening procedure consisted of: 1) apical 4-chamber (A4C) image preprocessing and segmentation; 2) A4C left atrial and left ventricular chamber value extraction; and 3) dynamic intrabeat linear data classification by a one-dimensional convolutional neural network.

in an independent external data set using patients with symptomatic COPD as a control group.

Recently, fully automated echocardiographic interpretation, which can automatically measure cardiac parameters (ejection fraction, LV mass index, and left atrial volume index), has been widely used to assist in diagnosis (7,8,21). For the diagnosis of HFpEF, several studies have collected all the speckle-tracking echocardiographic parameters and diastolic indexes for phenotypic classification (23-25) or used AI machine learning to classify patient information and parameters (26,27). These studies achieved precise results, but the required collection of comprehensive patient information for further analysis was very time consuming. Furthermore, as accurate interpretation of echocardiographic images may rely largely on clinician experience, interpretation using trained AI models may be superior to conventional interpretation by inexperienced clinicians. In contrast, our system is fast, accurate, and based only on a single A4C view, and it allows automatic diagnostic work flow; thus, it has the potential to be a powerful prescreening tool for HFpEF diagnosis (**Central Illustration**). Taken collectively, our CNN model can provide faster and fully automated echocardiographic framework to facilitate diagnostic interpretation for HFpEF diagnosis rather than producing many echocardiographic parameters, which still require expert interpretation.

STUDY LIMITATIONS. With respect to clinical limitations, first, this model was built specifically for the diagnosis of HFpEF and therefore did not include information on HF with reduced ejection fraction. Second, ultrasound artifacts caused by confounding factors, including those related to cauterization, artificial valves, and ventricular assist devices, could affect image quality and image segmentation results and eventually lead to inaccuracy in the linear data classification (28).

In terms of study limitations, first, we had fewer data on the HFpEF group than on the control group. Increasing the amount of training data might improve the moderate association of AI-estimated results and the export measurements. Second, each Digital Imaging and Communications in Medicine file must comprise 60 frames to ensure the inclusion of at least 1 complete systolic-diastolic cycle. However, this may preclude the application of our AI model in cases of extreme bradycardia. Also, in our AI model, only 1 echocardiographic view was used for HFpEF prescreening. We believe that combining multiview echocardiography could yield improved predictions. Third, concerning image quality, the chamber boundaries were blurred, and differences in machine vendors may affect AI detection accuracy. Fourth, as our data set contained only HFpEF data, additional data sets for training would be necessary for other types of cardiac diseases. Fifth, we used a fixed number of ultrasound frames to ensure that the data contained at least 1 complete cardiac cycle to extract the intrabeat dynamics. However, depending on the subject's heart rate, the data could contain multiple heartbeats, and the role of interbeat dynamics will need further investigation.

CONCLUSIONS

A fast and precise AI HFpEF prescreening tool based on tracking intrabeat dynamic changes in LV and left atrial length, width, area, and volume in the A4C view was developed. The use of LV intrabeat dynamic area changes had the highest accuracy (91%) in detecting HFpEF, with an area under the curve of 0.95. Our AI model can be a reliable prescreening tool compared with the time-consuming conventional HFpEF diagnosis process; moreover, it can provide quantitative information to assist inexperienced observers in the diagnosis HFpEF.

FUNDING SUPPORT AND AUTHOR DISCLOSURES

This work was supported by the Higher Education Sprout Project of National Chiao Tung University, Department of Medical Research, MacKay Memorial Hospital (Taipei City), and the Ministry of Education of Taiwan. The manuscript was edited by Wallace Academic Editing. The authors have reported that they have no relationships relevant to the contents of this paper to disclose.

ADDRESS FOR CORRESPONDENCE: Dr Chung-Lieh Hung, Department of Cardiology, Mackay Memorial Hospital, Taipei, Taiwan. E-mail: jotaro3791@gmail.com. OR Dr Shien-Fong Lin, Institute of Biomedical Engineering, College of Electrical and Computer Engineering, National Chiao Tung University, 1001 University Road, Hsinchu 30010, Taiwan. E-mail: linsf5402@nctu.edu.tw.

PERSPECTIVES

COMPETENCY IN MEDICAL KNOWLEDGE: This AI-assisted system provides fast (<3 min for each case) and accurate prescreening for HFpEF in a single A4C view on the basis of intrabeat dynamics changes before diagnostic confirmation using conventional echocardiography and blood tests.

COMPETENCY IN PATIENT CARE AND PROCEDURAL SKILLS: The diagnostic criteria for HFpEF follow the 2016 guidelines of the American College of Cardiology and the American Heart Association and take into account the diagnostic algorithm of the European Society of Cardiology: ejection fraction ≥50%, LV mass index ≥115 mL/m^2 in men and ≥95 mL/m^2 in women, left atrial volume index ≥34 mL/m^2, tissue Doppler–derived early myocardial relaxation (e′) < 9 cm/s, and LV passive filling (E/e′) >13, and a BNP concentration > 35 pg/mL. Our results demonstrate that AI can provide rapid prescreening to assist in diagnosis prior to conventional procedures such as echocardiography and can be of particular help to inexperienced clinicians.

TRANSLATIONAL OUTLOOK: This study demonstrates that our AI system has high accuracy in prescreening of HFpEF. However, collecting more A4C view data from patients with HFpEF is necessary for further system improvement.

REFERENCES

1. Lang RM, Badano LP, Mor-Avi V, et al. Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. J Am Soc Echocardiogr 2015;28:1–39.e14.

2. Tromp J, Teng TH, Tay WT, et al. Heart failure with preserved ejection fraction in Asia. Eur J Heart Fail 2019;21:23–36.

3. Sharma K, Kass DA. Heart failure with preserved ejection fraction: mechanisms, clinical features, and therapies. Circ Res 2014;115:79–96.

4. Ponikowski P, Voors AA, Anker SD, et al. 2016 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure: the Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure of the European Society of Cardiology (ESC) developed with the special contribution of the Heart Failure Association (HFA) of the ESC. Eur Heart J 2016;37:2129–200.

5. Pieske B, Tschope C, de Boer RA, et al. How to diagnose heart failure with preserved ejection fraction: the HFA-PEFF diagnostic algorithm: a consensus recommendation from the Heart Failure Association (HFA) of the European Society of Cardiology (ESC). Eur Heart J 2019;40:3297–317.

6. Sengupta PP, Shrestha S, Berthon B, et al. Proposed requirements for cardiovascular imaging-related machine learning evaluation (PRIME): a checklist: reviewed by the American College of Cardiology Healthcare Innovation Council. J Am Coll Cardiol Img 2020;13:2017–35.

7. Ruijsink B, Puyol-Anton E, Oksuz I, et al. Fully automated, quality-controlled cardiac analysis from CMR: validation and large-scale application to characterize cardiac function. J Am Coll Cardiol Img 2020;13:684–95.

8. Zhang J, Gajjala S, Agrawal P, et al. Fully automated echocardiogram interpretation in clinical practice. Circulation 2018;138:1623–35.

9. Roberts E, Ludman AJ, Dworzynski K, et al. The diagnostic accuracy of the natriuretic peptides in heart failure: systematic review and diagnostic meta-analysis in the acute care setting. BMJ 2015;350:h910.

10. Steinberg BA, Zhao X, Heidenreich PA, et al. Trends in patients hospitalized with heart failure and preserved left ventricular ejection fraction: prevalence, therapies, and outcomes. Circulation 2012;126:65–75.

11. Pfeffer MA, Pitt B, McKinlay SM. Spironolactone for heart failure with preserved ejection fraction. N Engl J Med 2014;371:181–2.

12. Liao JN, Chao TF, Kuo JY, et al. Age, sex, and blood pressure-related influences on reference values of left atrial deformation and mechanics from a large-scale asian population. Circ Cardiovasc Imaging 2017;10:e006077.

13. Nagueh SF, Smiseth OA, Appleton CP, et al. Recommendations for the evaluation of left ventricular diastolic function by echocardiography: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. J Am Soc Echocardiogr 2016;29:277–314.

14. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. Lect Notes Comput Sci 2015;9351:234–41.

15. Maier A, Syben C, Lasser T, Riess C. A gentle introduction to deep learning in medical image processing. Z Med Phys 2019;29:86–101.

16. Baskaran L, Maliakal G, Al'Aref SJ, et al. Identification and quantification of cardiovascular structures from CCTA: an end-to-end, rapid, pixel-wise, deep-learning method. J Am Coll Cardiol Img 2020;13:1163–71.

17. Hannun AY, Rajpurkar P, Haghpanahi M, et al. Cardiologist-level arrhythmia detection and

classification in ambulatory electrocardiograms using a deep neural network. Nat Med 2019;25:65–9.

18. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837–45.

19. Sasaki Y. The truth of the F-measure. Available at: https://www.cs.odu.edu/~mukka/cs795sum09dm/Lecturenotes/Day3/F-measure-YS-26Oct07.pdf. Accessed May 17, 2021.

20. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22:276–82.

21. Kusunose K, Abe T, Haga A, et al. A deep learning approach for assessment of regional wall motion abnormality from echocardiographic images. J Am Coll Cardiol Img 2020;13:374–81.

22. Bianco CM, Farjo PD, Ghaffar YA, Sengupta PP. Myocardial mechanics in patients with normal LVEF and diastolic dysfunction. J Am Coll Cardiol Img 2020;13:258–71.

23. Shah SJ, Katz DH, Selvaraj S, et al. Phenomapping for novel classification of heart failure with preserved ejection fraction. Circulation 2015;131:269–79.

24. Omar AMS, Narula S, Abdel Rahman MA, et al. Precision phenotyping in heart failure and pattern clustering of ultrasound data for the assessment of diastolic dysfunction. J Am Coll Cardiol Img 2017;10:1291–303.

25. Katz DH, Deo RC, Aguilar FG, et al. Phenomapping for the identification of hypertensive patients with the myocardial substrate for heart failure with preserved ejection fraction. J Cardiovasc Transl Res 2017;10:275–84.

26. Ahmad T, Lund LH, Rao P, et al. Machine learning methods improve prognostication, identify clinically distinct phenotypes, and detect heterogeneity in response to therapy in a large cohort of heart failure patients. J Am Heart Assoc 2018;7:e008081.

27. Angraal S, Mortazavi BJ, Gupta A, et al. Machine learning prediction of mortality and hospitalization in heart failure with preserved ejection fraction. J Am Coll Cardiol HF 2020;8:12–21.

28. Quien MM, Saric M. Ultrasound imaging artifacts: how to recognize them and how to avoid them. Echocardiography 2018;35:1388–401.

KEY WORDS artificial intelligence, chronic obstructive pulmonary disease, convolutional neural network, heart failure with preserved ejection fraction, U-net

APPENDIX For supplemental methods, figures, and tables, please see the online version of this paper.